Amendments to the Specification:

Paragraph beginning on page 1, line 10

Currently, there is a growing trend to converge voice and data networks so that both utilize the same network infrastructure. The currently available systems that combine voice and data have limited applications and scope. An example is Automatic Call Distribution (ACD), which permits service agents in call centers to access customer ~~filed~~ files in conjunction with incoming telephone calls. ACD centers, however, remain costly and difficult to deploy, requiring custom systems integration in most cases. Another example is the voice logging/auditing system used by emergency call centers (e.g., 911) and financial institutions. Deployment has been limited due to the limited scalability of the system since voice is on one network and data is on another, both tied together by awkward database linkages.

Paragraph beginning on page 2, line 7

An increasingly common IP telephony paradigm consists of telephone and data tightly coupled on IP ~~packet based~~ packet-based, switched, multimedia networks where voice and data share a common transport mechanism. It is expected that this paradigm will spur the development of a wealth of new applications that take advantage of the simultaneous delivery of voice and data over a single unified fabric.

Paragraph beginning on page 3, line 27

The audio codec 54 functions to encode audio signals from the audio source (e.g., microphone) for transmission over the network and to decode the received audio data for output to a loudspeaker. All H.323 audio terminals must be capable of encoding and decoding speech in accordance with G.711 including both A-law and [[i]]µ-law encoding. Other types of audio that may be supported include G.722, G.723, G.728 and G.729.

Paragraph beginning on page 3, line 32

The data channel supports telematic ~~application~~ applications such as electronic whiteboards, still image transfer, file exchange, database access, real time audiographics conferencing (T.120), etc. The system control unit 56 provides services as defined in the H.245 and H.225.0 standards. For example, the system control unit provides signaling for proper operation of the H.323 terminal, call control, capability exchange, signaling of commands and indications and messaging to describe

the content of logical channels. The H.225.0 Layer 64 is operative to format the transmitted video, audio, data and control streams into messages for output to the network interface. It also functions to retrieve the received video, audio, data and control ~~steams~~ streams from messages received from the network interface 68.

Paragraph beginning on page 4, line 7

The gateway functions to convert voice from the IP domain to the PSTN domain. In particular, it converts IP packetized voice to a format that can be accepted by the PSTN. The actual format depends ~~of~~ on the type of media and protocol used for connecting to the PSTN (e.g., T1, E1, ISDN BRI, ISDN PRI, analog lines, etc.). The gateway provides the appropriate translation between different video, audio and data transmission formats and between different communications procedures and medias.

Paragraph beginning on page 7, line 17

In order to combat the delay problems, many devices implement a jitter buffer on the receive side. If packets are only delayed on the network, arriving at the receiver before the jitter buffer underflows, the receive side will hear the sound as it was ~~original~~ originally transmitted by the local endpoint. If, however, packets are dropped or packets are delayed too much and the jitter buffer underflows (i.e. becomes empty), the receiving device either (1) replays the last packet received or (2) it injects a silence. Thus, in the event packets are dropped or are delayed excessively causing jitter buffer underflow, the sound that is played on the receive side is not the original sound that was transmitted.

Paragraph beginning on page 7, line 25

Many audio applications including voice require that the audio (or voice) be recorded, at one or both ends of a conversation. A block diagram illustrating a prior art ~~packet based~~ packet-based four channel audio recorder is shown in Figure 4. The system, generally referenced 80, comprises a packet network 88 to which are connected a plurality of endpoints 82, such as endpoints A and B. Each endpoint comprises a loudspeaker (not shown) for generating audio and a microphone for converting audio, i.e. voice, to an electrical signal. Each endpoint is operative to receive an Rx signal 90 from the other endpoint and to generate a Tx signal 92 to the other side.

Paragraph beginning on page 9, line 1

The present invention provides an apparatus for and a method of audio recording in ~~packet based~~ packet-based telephony systems. Using the present invention, the equivalent of four audio channels are recorded utilizing only two recording channels. Each channel recorded comprises the stream of packets (e.g., RTP packets) generated and transmitted by each endpoint to the other side. The RTP packets include the samples generated by the particular endpoint in addition to an indication (e.g., a pointer) of the samples received from the other side actually played by the endpoint. Note that the audio played on an endpoint is not necessarily the samples received from the other side.

Paragraph beginning on page 9, line 30

Since the recorder receives ~~from each endpoint,~~ the audio signal that was generated and transmitted from each endpoint, it can reconstruct the audio signal that was actually played on the endpoint. To playback an audio signal, the recording device needs to know the samples that were actually played on each endpoint. The recorder is provided knowledge of the audio played on the other end via information transmitted in the data sample packets it receives. Each endpoint is adapted to include an indication of the audio that is played, with the packet of data samples sent to the recorder.

Paragraph beginning on page 13, line 10

Fig. 4 is a block diagram illustrating a prior art ~~packet based~~ packet-based four-channel audio recorder;

Paragraph beginning on page 16, line 13

A block diagram illustrating the structure of an endpoint in the packet network in more detail is shown in Figure 5. For clarity sake, endpoint A only is shown in more detail with the structure of endpoint B being identical. Each endpoint, generally referenced 130, comprises an input port 132, output port ~~150~~ 152, packet processor 134, jitter buffer 136, D/A converter 138, amplifier and analog interface circuitry 140, microphone 144, analog amplifier circuit 146 and A/D converter 148.

Paragraph beginning on page 17, line 26

As described above, in a prior art recording device this requires four audio channels, including the audio played and generated at both endpoint, to be transmitted to the recording device.

Using the recording system of the present invention, the four audio channels can be effectively recorded using only two ~~channel~~ channels including one channel transmitted from each endpoint.

Paragraph beginning on page 19, line 7

In a second embodiment, a single centralized recording device is not used. Each endpoint comprises means for storing the call details and recording the content of the RTP packet stream (including the header and header extension of the RTP packets) sent to the other side of the connection rather than sending the packets over a TCP/IP connection to a recording device. The recording device may be connected to the endpoint device by any suitable means such as RS-232, USB, IEEE 1394, other parallel or serial means, wireless means, optical means, etc. or can be a part of the endpoint itself (e.g., a flash integrated circuit or memory module on the endpoint board).

Paragraph beginning on page 22, line 27

A flow diagram illustrating the playback method of the present invention performed on the recording device is shown in Figure 10. This method is performed when the recording device is requested to play back the audio that was played on one of the endpoints. To play back the audio generated on an endpoint, the sample contents of the RTP packet in order of RTP packet time are retrieved and played back.

Paragraph beginning on page 25, line 1

In the event the RTP packets are compressed, the endpoints must be adapted to decompress ~~then~~ them before performing the method of the present invention. All references (i.e. pointers) are to uncompressed samples.